

E. Solobodyansky

RECEIVED
JUN 21 2001
TECH CENTER 1600/2900 1652

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/402,093A
DATE: 05/16/2001
TIME: 12:40:50

Input Set : A:\001560-373.ST25.txt
Output Set: N:\CRF3\05162001\I402093A.raw

ENTERED

4 <110> APPLICANT: Ohsuye, Kazuhiro
5 Yabuta, Masayuki
6 Suzuki, Yuji
8 <120> TITLE OF INVENTION: Process for Producing Peptides Using a Helper Peptide
10 <130> FILE REFERENCE: 001560-373
12 <140> CURRENT APPLICATION NUMBER: US 09/402,093A
13 <141> CURRENT FILING DATE: 1999-09-29
15 <150> PRIOR APPLICATION NUMBER: PCT/JP99/00406
16 <151> PRIOR FILING DATE: 1999-01-29
18 <150> PRIOR APPLICATION NUMBER: JP 10-32272
19 <151> PRIOR FILING DATE: 1998-01-30
21 <160> NUMBER OF SEQ ID NOS: 24
23 <170> SOFTWARE: PatentIn version 3.0
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 4
27 <212> TYPE: PRT
28 <213> ORGANISM: Artificial Sequence
30 <220> FEATURE:
31 <223> OTHER INFORMATION: Amino acid sequence adjacent to a site cleaved by enterokinase.
33 <400> SEQUENCE: 1
35 Asp Asp Asp Lys
36 1
38 <210> SEQ ID NO: 2
39 <211> LENGTH: 4
40 <212> TYPE: PRT
41 <213> ORGANISM: Artificial Sequence
43 <220> FEATURE:
44 <223> OTHER INFORMATION: Amino acid sequence adjacent to a site cleaved by blood
45 coagulation Factor Xa.
47 <400> SEQUENCE: 2
49 Ile Glu Gly Arg
50 1
52 <210> SEQ ID NO: 3
53 <211> LENGTH: 7
54 <212> TYPE: PRT
55 <213> ORGANISM: Artificial Sequence
57 <220> FEATURE:
58 <223> OTHER INFORMATION: Amino acid sequence containing a site cleaved by renin.
60 <400> SEQUENCE: 3
62 Pro Phe His Leu Leu Val Tyr
63 1 5
65 <210> SEQ ID NO: 4
66 <211> LENGTH: 6
67 <212> TYPE: PRT
68 <213> ORGANISM: Artificial Sequence
70 <220> FEATURE:
71 <223> OTHER INFORMATION: Amino acid sequence of helper peptide.

RAW SEQUENCE LISTING DATE: 05/16/2001
 PATENT APPLICATION: US/09/402,093A TIME: 12:40:50

Input Set : A:\001560-373.ST25.txt
 Output Set: N:\CRF3\05162001\I402093A.raw

```

73 <400> SEQUENCE: 4
75 Val Asp Asp Asp Asp Lys
76 1 5
78 <210> SEQ ID NO: 5
79 <211> LENGTH: 6
80 <212> TYPE: PRT
81 <213> ORGANISM: Artificial Sequence
83 <220> FEATURE:
84 <223> OTHER INFORMATION: Amino acid sequence of helper peptide.
86 <400> SEQUENCE: 5
88 Gly Cys His His His His
89 1 5
91 <210> SEQ ID NO: 6
92 <211> LENGTH: 10
93 <212> TYPE: PRT
94 <213> ORGANISM: Artificial Sequence
96 <220> FEATURE:
97 <223> OTHER INFORMATION: Amino acid sequence containing a chemically cleaved site.
99 <400> SEQUENCE: 6
101 Pro Gly Gly Arg Pro Ser Arg His Lys Arg
102 1 5 10
104 <210> SEQ ID NO: 7
105 <211> LENGTH: 10
106 <212> TYPE: PRT
107 <213> ORGANISM: Artificial Sequence
109 <220> FEATURE:
110 <223> OTHER INFORMATION: Amino acid sequence of helper peptide.
112 <400> SEQUENCE: 7
114 His Arg His Lys Arg Ser His His His His
115 1 5 10
117 <210> SEQ ID NO: 8
118 <211> LENGTH: 5
119 <212> TYPE: PRT
120 <213> ORGANISM: Artificial Sequence
122 <220> FEATURE:
123 <223> OTHER INFORMATION: Amino acid sequence containing a site cleaved by Kex2 protease.
125 <400> SEQUENCE: 8
127 Ser Asp His Lys Arg
128 1 5
130 <210> SEQ ID NO: 9
131 <211> LENGTH: 23
132 <212> TYPE: PRT
133 <213> ORGANISM: Artificial Sequence
135 <220> FEATURE:
136 <223> OTHER INFORMATION: Amino acid sequence containing a position cleaved by OmpT.
138 <400> SEQUENCE: 9
140 Gln Met His Gly Tyr Asp Ala Glu Leu Arg Leu Tyr Arg Arg His His
141 1 5 10 15
143 Arg Trp Gly Arg Ser Gly Ser

```

RAW SEQUENCE LISTING

DATE: 05/16/2001

PATENT APPLICATION: US/09/402,093A

TIME: 12:40:50

Input Set : A:\001560-373.ST25.txt

Output Set: N:\CRF3\05162001\I402093A.raw

```

144          20
146 <210> SEQ ID NO: 10
147 <211> LENGTH: 20
148 <212> TYPE: PRT
149 <213> ORGANISM: Artificial Sequence
151 <220> FEATURE:
152 <223> OTHER INFORMATION: Amino acid sequence containing a position cleaved by OmpT.
154 <400> SEQUENCE: 10
156 Gln Met His Gly Tyr Asp Ala Glu Leu Arg Leu Tyr Arg Arg His His
157 1          5          10          15
159 Gly Ser Gly Ser
160          20
162 <210> SEQ ID NO: 11
163 <211> LENGTH: 69
164 <212> TYPE: DNA
165 <213> ORGANISM: Artificial Sequence
167 <220> FEATURE:
168 <223> OTHER INFORMATION: Nucleotide sequence coding for an amino acid sequence containing
169      a site cleaved by OmpT
171 <220> FEATURE:
172 <221> NAME/KEY: CDS
173 <222> LOCATION: (1)..(69)
175 <400> SEQUENCE: 11
176 cag atg cat ggt tat gac gcg gag ctg cgg ctg tat cgc cgt cat cac          48
177 Gln Met His Gly Tyr Asp Ala Glu Leu Arg Leu Tyr Arg Arg His His
178 1          5          10          15
180 cgg tgg ggt cgt tcc gga tcc          69
181 Arg Trp Gly Arg Ser Gly Ser
182          20
185 <210> SEQ ID NO: 12
186 <211> LENGTH: 23
187 <212> TYPE: PRT
188 <213> ORGANISM: Artificial Sequence
190 <220> FEATURE:
191 <223> OTHER INFORMATION: Amino acid sequence containing a site cleaved by OmpT.
193 <400> SEQUENCE: 12
195 Gln Met His Gly Tyr Asp Ala Glu Leu Arg Leu Tyr Arg Arg His His
196 1          5          10          15
199 Arg Trp Gly Arg Ser Gly Ser
200          20
203 <210> SEQ ID NO: 13
204 <211> LENGTH: 47
205 <212> TYPE: DNA
206 <213> ORGANISM: Artificial Sequence
208 <220> FEATURE:
209 <223> OTHER INFORMATION: Nucleotide sequence coding for an amino acid sequence containing
210      a site cleaved by OmpT
212 <400> SEQUENCE: 13
213 tggttatgac gcggagctcc gcctgtatcg ccgtcatcac ggttccg          47

```

RAW SEQUENCE LISTING

DATE: 05/16/2001

PATENT APPLICATION: US/09/402,093A

TIME: 12:40:50

Input Set : A:\001560-373.ST25.txt

Output Set: N:\CRF3\05162001\I402093A.raw

```

216 <210> SEQ ID NO: 14
217 <211> LENGTH: 55
218 <212> TYPE: DNA
219 <213> ORGANISM: Artificial Sequence
221 <220> FEATURE:
222 <223> OTHER INFORMATION: Nucleotide sequence coding for an amino acid sequence containing
223     a site cleaved by OmpT
225 <400> SEQUENCE: 14
226 gatccggaac cgtgatgacg gcgatacagg cggagctccg cgtcataacc atgca           55
229 <210> SEQ ID NO: 15
230 <211> LENGTH: 24
231 <212> TYPE: DNA
232 <213> ORGANISM: Artificial Sequence
234 <220> FEATURE:
235 <223> OTHER INFORMATION: Primer.
237 <400> SEQUENCE: 15
238 gactcagatc ttctgaggc cgat           24
241 <210> SEQ ID NO: 16
242 <211> LENGTH: 36
243 <212> TYPE: DNA
244 <213> ORGANISM: Artificial Sequence
246 <220> FEATURE:
247 <223> OTHER INFORMATION: Primer.
249 <400> SEQUENCE: 16
250 aaaggtacct tccgcatgcc gcggatgtcg agaagg           36
253 <210> SEQ ID NO: 17
254 <211> LENGTH: 20
255 <212> TYPE: DNA
256 <213> ORGANISM: Artificial Sequence
258 <220> FEATURE:
259 <223> OTHER INFORMATION: Primer.
261 <400> SEQUENCE: 17
262 aggccaggaa ccgtaaaaag           20
265 <210> SEQ ID NO: 18
266 <211> LENGTH: 29
267 <212> TYPE: DNA
268 <213> ORGANISM: Artificial Sequence
270 <220> FEATURE:
271 <223> OTHER INFORMATION: Primer.
273 <400> SEQUENCE: 18
274 aaaatgcatc gcatcgtaac cgtgcatct           29
277 <210> SEQ ID NO: 19
278 <211> LENGTH: 627
279 <212> TYPE: DNA
280 <213> ORGANISM: Artificial Sequence
282 <220> FEATURE:
283 <223> OTHER INFORMATION: Nucleotide sequence coding for a fusion protein comprising GLP-1,
284     helper peptide and beta-galactosidase protective peptide.
286 <220> FEATURE:

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/402,093A
 DATE: 05/16/2001
 TIME: 12:40:50

Input Set : A:\001560-373.ST25.txt
 Output Set: N:\CRF3\05162001\I402093A.raw

```

287 <221> NAME/KEY: CDS
288 <222> LOCATION: (82)..(543)
290 <400> SEQUENCE: 19
291 cccaggcttt acactttatg cttccggctc gtatgttggtg tggaattgtg agcggataac      60
293 aatttcacac aggaacacagc t atg acc atg att acg gat tca ctg gcc gtc      111
294                               Met Thr Met Ile Thr Asp Ser Leu Ala Val
295                               1           5           10
297 gtt tta caa cgt aaa gac tgg gat aac cct ggc gtt acc caa ctt aat      159
298 Val Leu Gln Arg Lys Asp Trp Asp Asn Pro Gly Val Thr Gln Leu Asn
299                               15           20           25
301 cgc ctt gca gca cat ccc cct ttc gcc agc tgg cgt aat agc gac gac      207
302 Arg Leu Ala Ala His Pro Pro Phe Ala Ser Trp Arg Asn Ser Asp Asp
303                               30           35           40
305 gcc cgc acc gat cgc cct tcc caa cag ttg cgc agc ctg aat ggc gaa      255
306 Ala Arg Thr Asp Arg Pro Ser Gln Gln Leu Arg Ser Leu Asn Gly Glu
307                               45           50           55
309 tgg cgc ttt gcc tgg ttt ccg gca cca gaa gcg gtg ccg gca agc ttg      303
310 Trp Arg Phe Ala Trp Phe Pro Ala Pro Glu Ala Val Pro Ala Ser Leu
311                               60           65           70
313 ctg gag tca gat ctt cct gac gcc gat act gtc gtc gtc ccc tca aac      351
314 Leu Glu Ser Asp Leu Pro Asp Ala Asp Thr Val Val Val Pro Ser Asn
315 75                               80           85           90
317 tgg cag atg cac ggt tac gat gcg atg cat ggt tat gac gcg gag ctc      399
318 Trp Gln Met His Gly Tyr Asp Ala Met His Gly Tyr Asp Ala Glu Leu
319                               95          100          105
321 cgc ctg tat cgc cgt cat cac ggt tcc gga tcc cct tct cga cat ccg      447
322 Arg Leu Tyr Arg Arg His His Gly Ser Gly Ser Pro Ser Arg His Pro
323                               110         115         120
325 cgg cat gcg gaa ggt acc ttt acc agc gat gtg agc tgg tat ctg gaa      495
326 Arg His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu
327                               125         130         135
329 ggt cag gcg gca aaa gaa ttc atc gcg tgg ctg gtg aaa ggc cgt ggt      543
330 Gly Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly
331                               140         145         150
333 taagtcgaca gccgcctaa tgagcgggct tttttttctc ggaattaatt ctcatgtttg      603
335 acagcttatc atcgataagc tttta      627
338 <210> SEQ ID NO: 20
339 <211> LENGTH: 154
340 <212> TYPE: PRT
341 <213> ORGANISM: Artificial Sequence
344 <220> FEATURE:
345 <223> OTHER INFORMATION: Amino acid sequence of a fusion protein comprising GLP-1,
346     helper peptide and beta-galactosidase protective peptide.
348 <400> SEQUENCE: 20
350 Met Thr Met Ile Thr Asp Ser Leu Ala Val Val Leu Gln Arg Lys Asp
351 1           5           10           15
354 Trp Asp Asn Pro Gly Val Thr Gln Leu Asn Arg Leu Ala Ala His Pro
355           20           25           30
358 Pro Phe Ala Ser Trp Arg Asn Ser Asp Ala Arg Thr Asp Arg Pro

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/402,093A

DATE: 05/16/2001

TIME: 12:40:51

Input Set : A:\001560-373.ST25.txt

Output Set: N:\CRF3\05162001\I402093A.raw